图書情報工作
第62卷 第17期 2018年9月

政府网站信息资源知识融合体系架构设计*

■ 张卫东　左娜　陆璐
吉林大学管理学院　长春 130022

摘要：［**目的/意义**］从知识融合的角度切入，解决政府网站信息资源的分散性、海量性与公众知识需求的完整性、精准化之间的矛盾，为用户提供更为有效的决策支持，提高政府网站信息资源的利用效率。［**方法/过程**］阐述政府网站信息资源的内涵及其价值需求，剖析知识融合的研究进展，在此基础上分析政府网站信息资源知识融合的流程，根据流程设计"数据级融合－概念级融合－决策级融合"的多层次的政府网站信息资源知识融合体系架构，并为每个具体层次构建相应的运行框架。［**结果/结论**］设计出的政府网站信息资源知识融合体系架构，为后续的扩展性与深入性研究提供理论支持与借鉴。

关键词：政府网站　信息资源　知识融合　知识地图
分类号：G254
DOI：10.13266/j.issn.0252－3116.2018.17.014

中国政府网数据显示，截至2017年12月1日，全国正在运行的政府网站已达到24 820家，如此规模的网站数量必然会产生大量的政府网站信息资源。以北京为例，2017年北京市各级政府主动公开政府信息共147.5万条，其中通过政府网站公开信息数665 500条，占总数的37.18%[1]。为了进一步了解北京市政府网站公开信息资源的利用状况，我们以热词"交通拥堵"在北京市政府网站进行检索，按相关度排序，共得出112 336条结果，并且可以看出各区治理实况、相关座谈会、调研情况等各个层面的信息交织在一起，为用户精准把握所需信息造成障碍。

可见，政府网站信息资源的爆炸式增长在为公众提供更丰富信息的同时，也呈现出冗余度高、分散性强和关联度低等制约性特点，这些制约性特点导致公众在海量的政府网站信息中选择有价值信息的能力被弱化，从而使他们产生选择困难焦虑和阅读筛选障碍，而"知识融合"理论的出现和不断发展，为政府部门解决此难题开辟了一条新的道路。知识融合的目的在于将大量动态、多源及分散的信息资源进行深度挖掘与组织，实现知识的再生、增值和创新，这一定位与政府网站信息资源组织与利用过程相一致，能进一步优化政府网站的知识服务效果。

1　政府网站信息资源的价值需求分析

1.1　政府网站信息资源概念

政府网站信息资源是指由各个政府部门在其相应的政府门户网站上发布的面向广大群众的数字资源，这些数字资源涵盖了经济、教育、管理和服务等各个方面，且具有公开性、时效性、权威性和易逝性等特点[2]。随着网络新媒体和移动新媒体的兴起与成熟，越来越多的政府部门将政府网站作为其发布政务信息的前沿阵地。

1.2　政府网站信息资源的价值需求

政府网站信息资源的价值是政府部门价值需求满足程度的主观感受，整个价值需求的过程与公众的信息知识需求直接对接。具体表现在以下3个方面：

（1）资源内容的广泛性需求。政府网站信息资源内容覆盖的广泛程度直接影响公众获取信息资源的广度。以中国公开信息整合服务平台为例，其信息资源的领域跨度较大，包括财政金融审计信息、港澳台侨工作信息、国防信息、城乡建设与环境保护信息、科技教育信息、民族宗教信息、卫生体育信息、综合政务信息等22个主题的信息，其中记录数最多的城乡建设与环境保护信息高达786 187条，记录数最少的国防信息也

* 本文系国家自然科学基金应急管理项目"政府网站信息资源多维语义知识融合研究"（项目编号：71740015）研究成果之一。
作者简介：张卫东（ORCID：0000－0001－7512－0814），副教授，博士，博士生导师，E-mail：wdzhang@jlu.edu.cn；左娜（ORCID：0000－0003－2449－7579），硕士研究生；陆璐（ORCID：0000－0001－5126－9892），硕士研究生。
收稿日期：2018－04－02　修回日期：2018－07－09　本文起止页码：112－119　本文责任编辑：刘远颖

达到了18 350条,广泛的信息资源是保证公民和政府满足基本信息资源价值需求的基础条件[3]。

(2)资源功能的多样化与精准化需求。这要求政府网站与移动终端及社交平台充分对接[4]。随着移动终端及社交媒体技术的不断成熟,移动政务也成为政府网站的考核指标之一,根据清华大学对全国415家政府门户网站的抽样调查,2017年59.2%的网站建设了政务微信,其中,12.8%的政务微信更新频率较低,27.2%尚未建立政务微信;28.4%的网站建设了APP且更新状况较好,22.2%的APP更新频率较低,49.4%尚未建设APP[5]。政府网站接入移动终端和社交平台,保证了资源功能的多样化与精准化,是政府网站信息资源建设的又一价值需求。

(3)资源价值的综合需求性。这种综合需求体现在可用性、更新性、关联性和实用性等各个方面,另外,综合需求性直接与公众获取政府网站信息资源的效率相关,因此,政府部门需要全面关联和整合服务资源,从而最大程度地满足公众的综合价值需求。

政府网站信息资源的价值需求决定了政府网站信息资源数量庞大、范围广泛、信息冗杂等特点,因此,借助知识融合手段来形成符合用户需求、基于逻辑关系的知识产品刻不容缓。

2 知识融合及其研究进展

"知识融合"这一概念源自于知识工程领域,经历了数据融合到信息融合再到知识融合的过程,于20世纪90年代正式得到相关学者的关注和研究[6],图书情报领域内的知识融合研究是在知识科学的视角下进行的[7]。当前国内外对知识融合的内涵并没有统一界定,本文在结合多位学者观点的基础上,总结认为:当前阶段的知识融合是一种关于信息资源组织和整合的研究范式,是一种技术工具和思维模式,知识融合针对的对象是动态、多源和分散的知识,运用语义网、关联数据和数据挖掘等手段对对象知识进行抽取、匹配和集成处理,弱化甚至消除对象知识的动态性、多源性和分散性的特点,最终形成新的知识层,实现知识增值。通过文献阅读,本文将学术界关于知识融合的研究方向总结为由外而内的4个层次的研究,即大数据环境层次、融合系统框架层次、关键技术层次、应用实践层次,如图1所示。除此之外,本文还对政府网站领域内的知识融合研究进展进行了梳理。

2.1 知识融合与大数据环境的联合探究

将知识融合置于大数据环境中进行研究是时代潮流所需。近年来,学术界相关学者根据大数据环境下信息资源数量庞大、结构复杂等特点[8],将知识融合看作大数据时代知识服务的增长点[9],并认为大数据环境下的知识融合环节发生了一定的新变化[10],这些新变化的出现促进了一些面向大数据环境的知识融合框架的重构。譬如,范昊将知识融合的过程模型应用于大数据知识服务框架中,把知识融合、知识服务和大数据三者有效地结合起来,不仅较好地满足了用户个性化、创新型的需求,更为相关研究提供了一定的理论支撑[11]。王曰芬则从知识生态重构的角度对大数据时代的知识融合关键问题进行了分析,并构建了DIKW价值链视角下的面向知识生态重构和学科创新服务的知识融合流程与框架[12]。

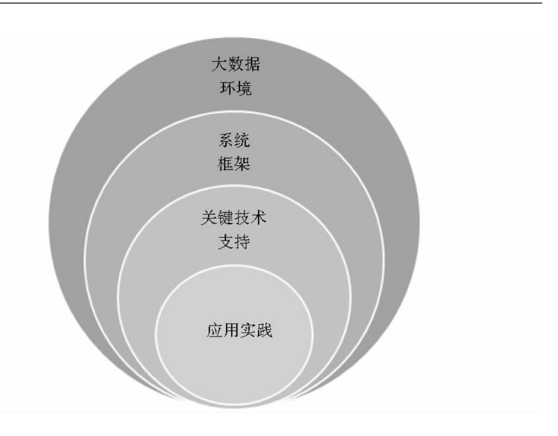


图1 学术界知识融合研究的4个层次

2.2 知识融合的系统框架构建研究

知识融合框架和知识融合系统的构建为知识融合研究提供了逻辑起点和统一支持。在知识融合框架的构建方面,房小可以开放世界的视角审视知识融合研究,为突破多源词表存在的异构性和知识扩展局限,她在波普尔世界理论的基础上论证了知识融合范式,并提出面向多源词表融合的框架体系"MtFFO"[13]。在知识融合系统方面,最具代表性的是P. M. Gray团队建设的KRAFT(Knowledge Reuse and Fusion Transform)系统,该系统开辟了知识融合系统架构的先河,为之后相关系统的建设奠定了强有力的基础[14]。在此之后,A. Nikolov等人建立了KnoFUSS系统[15],该知识融合系统结合了应用问题解决方法技术,根据不同的领域和任务,选择不同的方法解决问题。

2.3 知识融合的关键技术解析

知识融合的关键技术指的是实现知识融合的一系列算法研究,主要集中在语义规则、主题图、模糊集理论、贝叶斯网络和数据挖掘等方面。例如,缑锦设计了一种基于本体论和语义规则的知识融合算法,提高了

知识融合结果的准确性[16]；鲁慧民基于全信息理论，对传统主题图模式进行了元素间结构的扩展，并设计了一种面向多源异构知识融合的扩展主题图相似性算法[17]；韩立岩将信息融合处理技术移植到知识融合处理中，形成基于模糊集理论的知识融合算法，并得出了相关的处理流程和基于 Petri 网的融合模型[18]；E. Santos 等人将概率模型与传统意义上的贝叶斯网络相结合，使概率模型成功转化为贝叶斯知识基础，最终提出了贝叶斯知识融合算法[19]。

2.4 知识融合的应用实践探讨

近年来，知识融合的应用实践研究也如雨后春笋般出现，且涉及的领域十分广泛，包括图书情报与档案管理领域、电子制造领域、军事领域、计算机科学领域等。譬如，唐晓波将知识融合应用于个性化商品推荐中，从消费价值的角度深度分析了商品特征，挖掘出用户偏好，利用 DS 证据理论对传统协同过滤和知识过滤推荐算法进行融合，实现了个性化推荐[20]；A. Smirnov 等人利用分布式知识的网络方法，对知识融合、多 Agent 结构等进行分析，完成了知识融合技术在电子制造领域的应用[21]；沈旺将知识融合引入数字参考咨询服务，以领域本体映射技术为依托抽取、清洗和匹配知识本体库，并对所形成的元知识集进行知识融合，在此基础上完成数字参考咨询知识融合框架的构建，增强了数字参考咨询服务能力[22]；成全通过引入“知识因子”的概念，根据网络社区成员协作的不同实现形式提出以促进专家知识创新为目标的基于协同标注的知识融合实现机制[23]。

2.5 政府网站知识融合研究进展

虽然对政府网站的信息资源进行知识融合研究具备一定的合理性和科学性，但是就目前而言，学术界对该方面的研究还未形成较为成熟的体系，学者们往往侧重于运用知识融合某一阶段的理论或技术来解决政府网站不同对象信息资源的某些融合问题。例如，赵慧针对政府网站信息资源的“内孤岛化”问题，运用信息构建理论设计了政府网站信息资源的评价指标体系，旨在将碎片化的处于混沌的政府网站信息资源进行逻辑上的梳理，从而使政府网站信息资源呈现出一套结构较为清晰、内容较为集成的组织体系，这是知识采集及知识转换等过程的体现[24]；丁楠从关联数据技术的角度入手分析了政府信息的聚合问题，政府网站作为政府信息的最广泛的发布平台，其信息资源的聚合问题同样符合关联数据技术角度的研究，通过关联数据技术将政府外部信息和内部数据整合到一起，有利于政府信息资源的聚合，有利于用户和政府进行科学决策，这是知识关联等过程的体现[25]；胡海波运用知识融合理论，提出了基于知识融合的电子政务知识服务模式，旨在解决政府的知识服务问题，政府网站作为电子政务建设的重要组成部分，该知识服务模式同样可以适用于政府网站信息资源的知识融合，该模式体现了知识抽取、知识转换、知识融合和知识服务等环节[26]；毛秀梅以政府应对网络舆情为背景，分析了知识采集、知识分析与表示和知识关联等技术，结合政府网络舆情知识服务流程，构建了基于 OGSA 的网络舆情知识服务架构，这对于政府网站信息资源的知识融合研究具有一定的借鉴意义[27]。本文在结合了前人研究的基础之上，着眼于政府网站信息资源知识融合的的全阶段，设计出多维度的知识融合框架，并提出了相应的知识融合的实现方法。

3 政府网站信息资源知识融合流程分析

了解和分析政府网站信息资源知识融合的流程，对于把握知识融合的体系架构具有很强的指导性作用，本文以政府网站资源本身所具有的特性为出发点，对其知识融合的关键流程进行了针对性研究。政府网站信息资源的知识融合流程包括 4 个重要的核心过程：知识采集、知识转换、知识融合和知识服务[28]。另外，本文还从知识的序化、重组和聚类程度的角度，将整个流程中的知识划分为 3 个不同的阶段，即广义知识、分类知识和关联聚类知识。政府网站信息资源知识融合流程如图 2 所示：

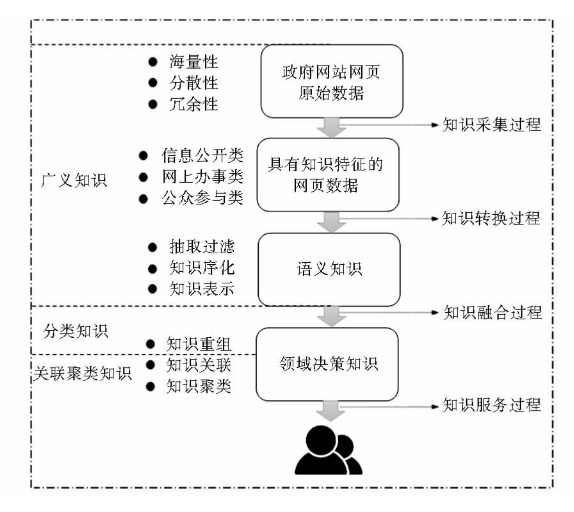


图 2 政府网站信息资源知识融合流程

3.1 知识采集

原始数据是整个知识融合流程得以运行的基本单元，对于具有海量性、分散性、冗余性等特点的政府网

站资源来说，其网页所含的原始数据信息同样处于动态、多样、繁杂和分布式的状态，因此流程中的第一个步骤是对处于这些状态的网站网页原始数据进行采集。政府网站信息资源包含信息公开、网上办事和公众参与等版块，主要涵盖了政策法规、机构简介、在线咨询、公众论坛、特色介绍、统计年鉴、便民服务等多个具体内容，在网页原始数据的采集中可以根据某一主题的要求对原始网页信息进行初步的、粗略的甄别，以筛选出具有知识特征的网页数据。

3.2 知识转换

知识采集过程已经对原始的网页数据作了初步甄别，并从中得到了具有知识特征的网页数据，知识转换过程的目标就是将这些具有知识特征的数据转换为知识资源，来实现网页资源内容的知识化[29]。知识转换的手段包括对采集数据的知识抽取、知识过滤、知识分解、知识序化和知识表示等。其中，经过简单的知识抽取、过滤和分解的资源还是处于一个杂乱无章的状态，而接下来的知识序化工作便通过对其进行统一的资源内容描述变无序为有序，最后通过明确的知识资源表示，成功将最初的具有知识特征的网页数据转换为具有逻辑性的语义知识。

3.3 知识融合与知识服务

知识融合算法是知识融合的关键支撑因素[30]。知识融合就是将具有逻辑性的语义知识通过算法使其升级为具有决策支持功能的领域知识。在这一过程中，需要运用一定的知识融合算法对前3个过程中获取与处理过的知识进行知识重组、知识分类和知识聚类等运算。通过知识重组对知识资源进行深度重组归类与合并，再通过知识关联将重组过的知识进行关联分析，挖掘隐藏在知识资源中的序化的相互关系，最后，通过知识聚类将主题相近或存在一定关系的知识进行聚集，从而为知识服务奠定坚实的基础。知识服务是整个流程的最后一个环节，主要目的就是将知识融合完成的知识提供给利用者，从而为其提供决策支持。

4 政府网站信息资源知识融合体系架构设计

本文在结合了KRAFT[14]、Knofuss[15]、SemFus[31]等较为成熟的知识融合系统的基础之上，从广义知识、分类知识和关联聚类知识的角度切入，构建了面向广义知识的数据级融合层，面向分类知识的概念级融合层和面向关联聚类知识的决策级融合层的3层体系[32]，如图3所示：

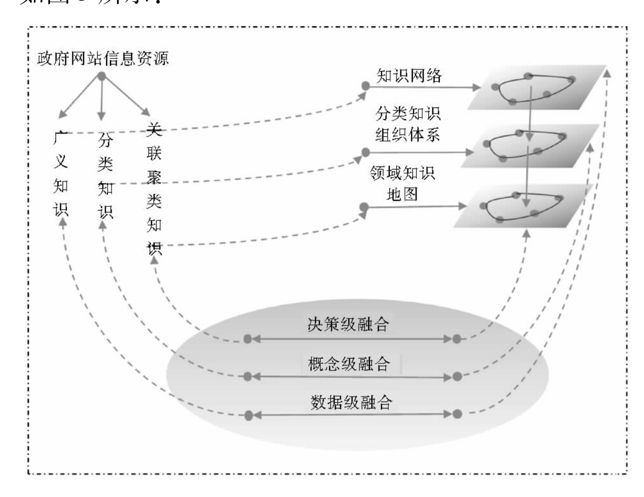


图3 政府网站信息资源知识融合体系架构

如第3节所述，政府网站信息资源在知识融合过程中会形成广义知识、分类知识和关联聚类知识，其中，广义知识形成于知识采集和知识转化的过程中，分类知识则产生于知识重组环节，而关联聚类知识则在知识关联和知识聚类环节中产生。广义知识是立足于内容层面的宏观性的概念，是指对经过语义化技术分析的政府网站信息资源的普遍性属性进行表征后的知识；分类知识是着眼于结构层面的知识，是指通过领域本体概念合并后的结构精简的、优化的知识；关联聚类知识则是致力于应用层面的知识，是指对广义知识和聚类知识的关联和聚类，是面向决策的完整的领域知识。

数据级融合是面向广义知识的融合，在这一融合过程中，借助语义描述等技术，构造生成知识网络，该知识网络的作用是对特定主题的资源进行宏观性的描述；概念级融合是面向分类知识的融合，它以数据级融合中形成的知识网络为基础，运用本体合并、领域概念获取等技术对知识网络进行划分，形成一定的分类组织体系结构；决策级融合是面向关联聚类知识的融合，它是以数据级融合中形成的知识网络和概念级融合中形成的分类组织体系为依托，通过知识发现技术、知识融合算法等方法，最终形成知识地图，为利用者提供其决策所需的知识支持。

4.1 面向广义知识的数据级融合

面向广义知识的数据级融合是政府网站信息资源知识融合体系架构的第一个层级，也是概念级融合和决策级融合的基础和运行条件。数据级融合的主要任务是利用RDF链接、数据关联等方式实现知识资源间初步的、简单的关联融合，最终建立起面向特定主题的

政府网站信息资源知识网络[33]。其框架如图 4 所示：

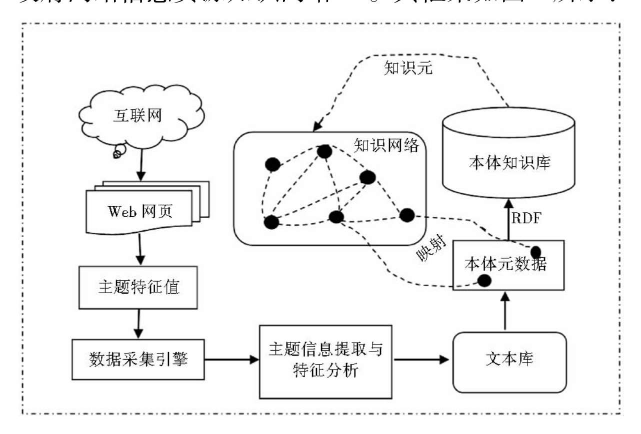


图 4 面向广义知识的数据级融合框架

图 4 反映的是面向广义知识的数据级融合的框架结构，在这一层级中，主要包括政府网站信息资源采集与预处理、本体建模、知识表示和知识网络构建等内容，具体流程如下：

（1）资源采集与预处理。存在于互联网中的政府网站 Web 网页中有大量的、无序的信息，这些信息构成了资源采集与预处理的资源来源，在获取这些资源后，需要根据实际需求对这些网页资源的主题特征值进行主题分类，并通过灵活使用自动化深度标引、页面去噪和网页切割等技术，来完成数据采集引擎所返回数据的主题信息提取和特征分析工作，然后将分析好的信息资源存入文本库。

（2）建立本体知识库。经过采集与预处理后的信息资源被存放在文本库中，这时的信息资源虽具备一定程度的主题特征，但其还需进一步的序化，由于元数据自身的限制性因素导致其无法高效准确地描述文本库中的政府网站信息资源，此时，将本体技术引入到元数据管理中来，从 SQL、XML、RDF 数据源中定位、抽取和转换成本体元数据，并以 RDF 的形式存储于本体知识库中。

（3）构建知识网络。本体知识库中的以 RDF 形式存在的资源属于半结构化的知识资源，这时，可以引入“知识元”的概念，并结合本体映射技术来确定海量的政府网站信息资源之间的隐性关系，通过本体语义关联方法将本体知识库中半结构化的知识资源通过可以被理解的知识网络的形式呈现出来。

面向广义知识的数据级融合需要借助定题采集、语义描述技术以及知识元 - 本体映射技术等重要方法和关键技术来实现。

（1）定题采集。信息采集是政府网站信息资源知识融合的基础，政府网站的信息资源大多以主题的形式分布和呈现在网站页面中，针对其这一属性特征，本文采取定题采集的方法，利用机器学习的自适应定题采集等手段，研究政府网站主题词表，而后采用基于统计的网页主题信息提取方法，利用网页 DOM 树结构、页面去噪、汉语自动分词等途径实现主题信息的提取。

（2）语义描述技术。语义描述技术实现的是政府网站信息资源主题内容的形式化表达，即结构化语义功能的扩展。在利用政府网站信息资源元数据标准描述政府网站信息资源的基础之上，结合本体技术进行更加丰富的描述，形成本体驱动的元数据模型，而后采用七步法构建领域知识本体。

（3）知识元 - 本体映射技术。知识元 - 本体映射技术实现的是对领域知识本体表征的知识元素进行统一对象化表示，其首先要解决的问题是知识元的抽取，初步定义知识元的语义组织框架，而后利用知识元 - 本体映射技术，将知识元与其相对应的知识本体进行映射，建立语义链接。

4.2 面向分类知识的概念级融合

面向分类知识的概念级融合是政府网站资源知识融合体系架构的中间层次的融合，起着承上启下的作用。概念级融合的主要任务就是将知识网络中的领域概念关联融合，完成更细粒度的概念合并，形成分类知识组织体系，最终达到消除冗余的目的。其框架如图 5 所示：

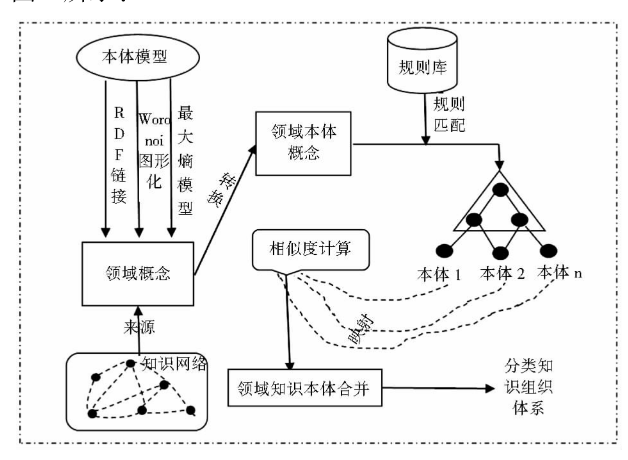


图 5 面向分类知识的概念级融合框架

在面向分类知识的概念级融合层中，主要包括政府网站信息资源领域概念获取、领域本体概念关系识别和领域知识本体合并等内容，具体流程如下：

（1）领域概念获取。领域概念的获取来源是知识网络中的资源，它包括实体概念、属性概念和关系概念 3 类[34]。领域概念可以通过利用 RDF 定义概念的功能，通过 RDF 链接、最大熵模型和 Voronoi 图形化的方

法来进行获取。

（2）领域本体概念关系识别。领域本体概念具备不同的表现形式，其关系类型包括上下位关系、实例关系、序列关系和定价关系等，因此，利用领域概念抽取算法，再结合基于融合规则的方法进行规则匹配，从而识别领域概念间关系，进而建立领域本体概念关系的层次结构，是构建分类知识组织体系必不可少的环节。

（3）领域知识本体合并。领域本体概念间的关系被识别后，则需要对这种关系进行下一步的应用，在领域概念关系的基础之上，通过映射的方法来构建领域知识本体映射模型，进而进行概念间语义关系的相似度计算，完成政府网站信息资源领域知识本体的合并，最终形成逻辑化的分类知识组织体系，减少知识冗余，提高用户的利用效率。

面向分类知识的概念级融合的实现需要依托于混合式领域概念获取法、自动抽取算法、领域知识本体合并等技术。

（1）混合式领域概念获取法。之所以称其为混合式，是因为此方法是基于统计的领域概念获取算法和基于规则的领域概念获取算法相结合的产物，本文将此法依据政府网站信息资源的特征进行改进，形成基于本体表示的获取算法，其原理如下：通过 RDF 链接进行定性推理，利用 Voronoi 图形式进行定量推理，使用最大熵模型进行权重计算，最后完成本体形式表示下的领域概念的获取工作。

（2）自动抽取算法。自动抽取算法主要实现的是对领域本体概念关系的识别与抽取，本文采用基于模式匹配的算法对领域资料文本的对应关系进行匹配，并采用基于统计的算法计算任意概念的统计度量，精准识别其间关系，将两种算法相结合，保证更为准确地识别和抽取领域概念关系。

（3）领域知识本体合并技术。对政府网站信息资源的领域本体进行合并可以达到减少知识网络中知识冗余的现象，领域知识本体的合并首先要建立领域本体概念分类体系；其次，建立并利用领域本体映射模型完成特征提取、选取概念对等步骤；最后，需要着重关注领域本体概念间语义的相似度计算。相似度计算的实现需要构建由本体解析模块、本体库和语义相似度计算等模块组成的运算框架。

4.3 面向关联聚类知识的决策级融合

面向关联聚类知识的决策级融合是政府网站信息资源知识融合体系架构中的最高层级，也是数据级融合和概念级融合价值的最终体现。这一层级的主要任务是利用一定的方法和算法将某一主题的政府网站信息资源中具有显性或隐性关系的分散的领域知识进行关联聚类，为用户提供蕴含决策支持功能的领域知识地图。其框架如图 6 所示：

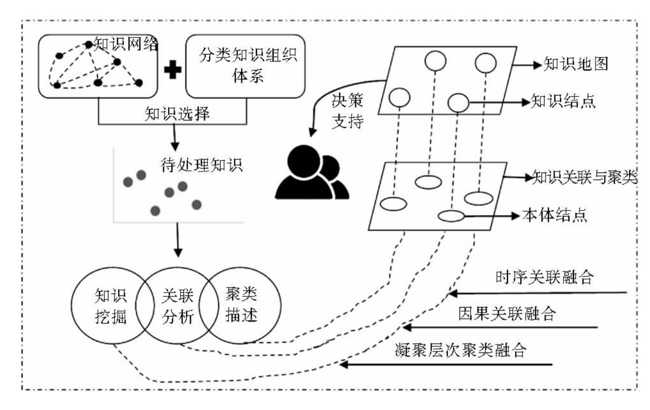


图 6　面向关联聚类知识的决策级融合框架

面向广义知识的数据级融合层中所形成的知识网络和面向分类知识的概念级融合层中所形成的分类知识组织体系中的某些领域知识与领域知识之间还存在着未被挖掘和利用的隐藏关系，这些领域知识将会被进行初步选择，形成待处理的知识。这些待处理的知识在面向关联聚类知识的决策级融合层中将会得到深入分析与处理，具体内容如下：

（1）知识关联聚类过程中的知识挖掘、关联分析和聚类描述应用。政府网站信息资源领域知识关联聚类的过程是指将相似性和关联性高的知识群组划分为相同群组，此过程的实现离不开知识挖掘、关联分析和聚类描述方法的应用，通过知识挖掘手段挖掘出存在一定关联关系的知识资源群组，利用关联分析将存在相互关系的知识资源群组的规律呈现出来，再利用聚类描述将主题相近的政府网站信息资源知识群组聚集在一起，形成多个知识资源类。

（2）基于关联挖掘和语义聚类的领域知识融合算法应用。知识挖掘、关联分析和聚类描述的方法可以发现和挖掘领域知识之间存在着的时序、因果以及凝聚层次关系，这些关系为领域知识的关联和聚类提供了可能。其中，基于时序关联的领域知识融合方法指按照时间序列的变化趋势之间的关系对领域知识进行融合，从而产生时序关联图。基于因果关联的领域知识融合方法是指遵循因果规律对领域知识进行融合，形成因果关联图。基于凝聚层次聚类的领域知识融合是指从知识聚类的角度出发对领域知识进行融合，形成聚类层次树。

（3）构建知识地图。本文通过提取关联聚类领域

知识中的知识点，并将知识点加以描述形成知识结点，再综合利用时序关联图、因果关联图和聚类层次树，完成领域知识地图的构建工作。知识地图是知识资源展示的图形导航工具，它可以为一定范围内的用户提供基于视图的知识获取、交流和开发利用服务[35]。最后面向关联聚类知识的决策级融合为解决用户问题提供了资源组织的基础，有利于更好地满足用户的知识需求，提高政府网站信息资源的开发与利用效率。

同样地，面向分类聚类知识的决策级融合的实现也离不开相应方法和关键技术的支撑，例如知识发现技术、知识融合算法以及可视化与人工交互技术的应用等。

（1）知识发现技术。知识发现技术主要实现了政府网站信息资源领域知识的关联聚类流程，通过面向文本内容和面向网络结构的知识挖掘技术，挖掘出政府网站中存在一定关联关系的知识资源，再利用知识关联技术，设定相应的关联规则，运用多重运算对挖掘出的知识资源进行关联分析，最后选定“聚类阈值”分解知识对象集合，运算“聚类近似值”，进行知识聚集，形成知识资源类。

（2）知识融合算法。本文主要采用时序关联融合算法、因果关联时序算法和凝聚层次聚类融合算法，分别从时间序列、因果关系、凝聚层次的角度实现政府网站信息资源的关联聚类知识的融合。

（3）可视化与人工交互技术。可视化与人工交互技术具备智能化的特点，可以更好的帮助用户享受知识融合的成果。例如，基于前文中的政府网站信息资源知识融合的脉络，建立知识地图或知识图谱，为用户提供立体、交互、可视化的决策支持，帮助用户直观地在短时间内获取所需的知识资源。

5 结语

在信息化浪潮的席卷下，政府网站信息资源的规模正以几何式的增长速度不断扩大，日益呈现出海量性、异构式、分散性等特点，在此背景下，政府网站信息资源的价值需求也在不断升级，如广泛性需求、多样化需求、精准化需求和综合性需求等。为了降低政府网站信息资源海量性等固有制约因素带来的影响并实现政府网站信息资源的价值需求，需要架构一套完整的应对体系。本文从“知识融合”的视角切入，在整合了前人相关研究的基础之上，从知识采集、知识转换、知识融合和知识服务4个过程分析了政府网站信息资源知识融合的流程，并从流程分析中划分出了领域知识的3个不同阶段，即广义知识阶段、分类知识阶段和关联聚类知识阶段，从3个阶段入手，完整设计了政府网站信息资源知识融合的体系架构，包括面向广义知识的数据级融合层、面向分类知识的概念级融合层和面向关联聚类知识的决策级融合层。该体系架构逻辑性强，具备一定的理论指导意义。本研究从宏观角度研究了政府网站信息资源的体系架构问题，但事实上，关于政府网站信息资源知识融合的建设也需要注重微观层面的研究，具体融合算法、实例验证以及测评等方面的研究将成为接下来的研究方向。

参考文献：

[1] 2017年北京市政府信息公开工作年度报告[EB/OL].[2018－03－27]. http://zhengwu. beijing. gov. cn/zwzt/2017nb/.

[2] 黄新平，王萍，李宗富. 政府网站原生数字政务信息资源馆藏建设研究[J]. 图书馆工作与研究，2017(6)：87－92.

[3] 中国政府公开信息整合服务平台[EB/OL].[2018－03－27]. http://govinfo. nlc. gov. cn/.

[4] 罗贤春，肖文芳，文庭孝. 政务信息资源价值需求分析：以政府门户网站为例[J]. 高校图书馆工作，2017，37(5)：47－51.

[5] 中国政府网[EB/OL].[2018－03－27]. http://www. gov. cn/.

[6] 刘晓娟，李广建，化柏林. 知识融合：概念辨析与界说[J]. 图书情报工作，2016，60(13)：13－19，32.

[7] 党洪莉. 知识科学视角下我国知识融合研究现状解析[J]. 情报杂志，2015，34(8)：158－162.

[8] 唐晓波，朱娟，杨丰华. 大数据环境下的知识融合框架模型研究[J]. 图书馆学研究，2016(1)：32－35，18.

[9] 唐晓波，魏巍. 知识融合：大数据时代知识服务的增长点[J]. 图书馆学研究，2015(5)：9－14，8.

[10] 张心源，邱均平. 大数据环境下的知识融合框架研究[J]. 图书馆学研究，2016(8)：66－70，43.

[11] 周利琴，范昊，潘建鹏. 基于知识融合过程的大数据知识服务框架研究[J]. 图书馆学研究，2017(21)：53－59.

[12] 王曰芬，岑咏华. 大数据时代知识融合体系架构设计研究[J]. 数字图书馆论坛，2016(10)：16－24.

[13] 严承希，房小可. 开放世界视角：面向多源词表的知识融合框架MtFFO研究[J]. 中国图书馆学报，2017，43(4)：114－129.

[14] PREECE A，HUI K，GRAY A，et al. The KRAFT architecture for knowledge fusion and transformation [J]. Knowledge based systems，2000，13（2）：113－120.

[15] NIKOLOV A，UREN V，MOTTA E. KnoFuss：a comprehensive architecture for knowledge fusion [C]//Proceedings of the 4th international conference on knowledge Capture. San Francisco：ACM，2007：185－186.

[16] 缑锦. 一种基于语义规则的知识融合方法[C]// 2006年全国理论计算机科学学术年会论文集. 重庆：中国计算机学会理论计算机科学专业委员会，2006：11.

[17] 鲁慧民，冯博琴，李旭. 面向多源知识融合的扩展主题图相似性

算法[J]. 西安交通大学学报,2010,44(2):20－24.
[18] 周芳,刘玉战,韩立岩. 基于模糊集理论的知识融合方法研究[J]. 北京理工大学学报(社会科学版),2013,15(3):67－73.
[19] SANTOS E. Fusing multiple bayesian knowledge sources[J]. International journal of approximate reasoning, 2011, 52(7):935－947.
[20] 朱娟,唐晓波. 基于三层知识融合模型的个性化商品推荐[J]. 图书馆学研究,2017(5):24－30.
[21] SMIRNOV A,PASHKIN M,CHILOV N,et al. Ksnet - approach to knowledge fusion distributed sources[J]. Computing and informatics,2012,22(2):105－142.
[22] 沈旺,李亚峰,侯昊辰. 数字参考咨询知识融合框架研究[J]. 图书情报工作,2013,57(19):139－143.
[23] 成全. 基于协同标注的科研社区知识融合机制研究[J]. 情报理论与实践,2011,34(8):20－25.
[24] 赵慧,刘君. 以用户为中心的信息构建与网络治理——信息构建理论视野下的政府网站信息资源“去孤岛化”研究[J]. 公共管理学报,2013,10(1):128－136,144.
[25] 丁楠,王钰,潘有能. 基于关联数据的政府信息聚合研究[J]. 情报理论与实践,2015,38(7):76－79,85.
[26] 胡海波,高鹏. 基于知识融合的电子政务知识服务模式及系统模型[J]. 情报工程,2018,4(1):31－41.
[27] 毛秀梅,杨晔. 面向政府的网络舆情知识服务体系构建研究[J]. 情报科学,2016,34(9):124－128.
[28] 徐绪堪,郑昌兴,蒋勋. 基于粒度原理的知识组织模型构建[J]. 图书与情报,2013(6):8－12,56.
[29] 祝振媛,李广建.“数据－信息－知识”整体视角下的知识融合初探——数据融合、信息融合、知识融合的关联与比较[J]. 情报理论与实践,2017,40(2):12－18.
[30] 周芳,王鹏波,韩立岩. 多源知识融合处理算法[J]. 北京航空航天大学学报,2013,39(1):109－114.
[31] NOUGHABI H A,KAHANI M,BEHKAMAL B. SemFus: semantic fusion framework based on JDL [C]//Innovations and advances in computer, information, systems sciences, and engineering. New York: Springer,2013: 583－594.
[32] 黄新平. 政府网站信息资源多维语义知识融合研究[D]. 长春:吉林大学,2017.
[33] 陈为东,王萍,王益成,等. 政府网站信息资源的多维语义知识融合结构体系及策略研究[J]. 情报理论与实践,2017,40(6):111－116.
[34] 陆佳莹,袁勤俭,黄奇,等. 基于概念格理论的产品领域本体构建研究[J]. 现代图书情报技术,2016(5):38－46.
[35] 胡昌平,张晶,陈果. 面向用户的知识地图分类构建与应用[J]. 数字图书馆论坛,2017(4):2－9.

作者贡献说明：
张卫东：提出论文研究框架、修改论文；
左娜：撰写论文；
陆璐：搜集资料及英文内容的翻译。

Knowledge Fusion System Architecture Design of Government Website Information Resources

Zhang Weidong Zuo Na Lu Lu
School of Management, Jilin University, Changchun 130022

Abstract: [**Purpose/significance**] From the perspective of knowledge fusion, this paper try to solve the contradiction between the scattered and massive information resources on government websites and the integrated and precise knowledge needs of the public, so that can provide decision support for users and improve the utilization efficiency of information resources on government websites. [**Method/process**] Firstly, this paper explained the connotation and value requirements of government website information resources, and reviewed the research progress of knowledge fusion. Then, it analyzed the knowledge fusion process of government website information resources. Finally, based on the process analysis, it designed a knowledge fusion system architecture of government website information, which involves several levels included “data level fusion-concept level fusion-decision level fusion”, and we also built appropriate operational framework for each specific level. [**Result/conclusion**] The knowledge fusion system architecture of government website information we designed provides theoretical and referential support for scalable and in-depth research subsequently.

Keywords: government website information resource knowledge fusion knowledge map